DOCUMENT RESUME

ED 335 393 TM 017 009

AUTHOR Lai, Morris K.
TITLE Field-Based Concerns about Fourth-Generation Evaluation Theory.
PUB DATE Apr 91
NOTE 9p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 3-7, 1991).
PUB TYPE Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Educational Assessment; *Educational Research; Elementary Secondary Education; *Evaluation Methods; Evaluation Problems; Field Studies; *Naturalistic Observation; *Research Methodology; *Theory Practice Relationship
IDENTIFIERS *Fourth Generation Evaluation Theory

ABSTRACT

Some aspects of fourth generation evaluation procedures that have been advocated by E. G. Guba and Y. S. Lincoln were examined empirically, with emphasis on areas where there have been discrepancies between theory and field-based experience. In fourth generation evaluation, the product of an evaluation is not a set of conclusions, recommendations, or value judgments, but rather an agenda for negotiation of claims, concerns, and issues. This approach is distinguished from the earlier three generations of evaluation which dealt with measurement, description, and judgment respectively. Approximately 10 evaluations of educational programs conducted each year by the Curriculum Research and Development Group of the University of Hawaii College of Education over the past 15 years provided data for this analysis. An example involving the evaluation of federally-funded projects is provided. The data highlight dilemmas in the attempt to provide fourth generation evaluation, beginning with the expressed desires of stakeholders for the sort of information earlier generations of evaluation provided. An internal inconsistency appears to exist in the views of Guba and Lincoln about the impossibility of generalizing from one situation to another, even as they generalize about scientific theory. Guba and Lincoln have offered a theoretical model that promises to enhance evaluation if it can be applied in the real world, although the benefits of earlier models cannot be denied. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Morris K. Lai

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED338693

Field-based Concerns About Fourth-Generation Evaluation Theory

BEST COPY AVAILABLE

Morris K. Lai

Curriculum Research and Development Group
College of Education, University of Hawai'i

4 April 1991

Paper presented at the annual meeting of the American Educational Research Association, Chicago

TM017009

Field-based Concerns About Fourth-Generation Evaluation Theory

Morris K. Lai
University of Hawai'i

The purpose of this study was to investigate empirically of some of the major aspects of "Fourth-generation evaluation" procedures that have been advocated by Guba and Lincoln (1989). I was particularly interested in areas where there might have been notable discrepancies between the theory and the field-based experience involving evaluations done under contract.

Perspective/theoretical framework

Guba and Lincoln (1989) have eloquently argued that it is now time to move to "fourth generation evaluation," in which the product of an evaluation is not a set of conclusions, recommendations, or value judgments, but rather an agenda for negotiation of claims, concerns, and issues. They advocate what they call a constructivist approach, which begins with the assumption that realities are not objectively "out there," but instead are constructed by people. While making their case, Guba and Lincoln proclaim, describe, and critique the first three generations of evaluation. Table 1 on page 2 provides a summary of these evaluation generations.

Having been impressed by the eloquence of the arguments for fourth-generation evaluation, the Evaluation Office of the Curriculum Research and Development Group (CRDG) of the University of Hawai'i's College of Education attempted to apply some of Guba and Lincoln's procedures to a number of their contracted evaluations. Several major problems arose in our attempts to carry out fourth-generation evaluations.

Data source

CRDG has been conducting external evaluation of various educational and other programs for the past 15 years. Recently there have been about 10 separate evaluation contracts each year with a total budget of about $300,000. Clients have included the state Department of Education, the Department of Human Services, the state Department of Health and the University of Hawaii's School of Medicine, cooperative learning, performance-based learning, early childhood education,

The First Three Generations of Evaluation According to Guba and Lincoln

	1st Generation: Measurement	2nd Generation: Description	3rd Generation: Judgment
Role of evaluator	technical expert	describer	judge
Key characteristics	tests to measure effectiveness	objectives; formative evaluation	standards; objective evaluator
Examples	IQ & achievement tests	Ralph Tyler; Eight-year Study	Stake, CIPP, Eisner, Scriven
Flaws (according to G & L)*	weak on non-human evaluands	objectives not necessarily valid	evaluator may be reluctant to judge

*According to Guba and Lincoln, all three generations have at least the following major flaws or defects: a tendency toward managerial, a failure to accommodate value pluralism, and overcommitment to the scientific paradigm of inquiry.

Fourth-Generation Evaluation According to Guba and Lincoln

Role of evaluator	†❶ human instrument and human data analyst; ❷ illustrator & historian; ❸ mediator of judgmental process; ❹ collaborator, learner, teacher, reality shaper, change agent, & others not yet manifested
Key characteristics	primarily qualitative methods; no causally inferential statistics
Examples	none known
Flaws (according to G & L)	will someday be shown to be inadequate

†The numbers in this section correspond to the four generations of evaluation.

Chapter 2 programs, bilingual education, language arts, HIV/AIDS education, drug-free schools, and special needs schools.

The evaluations conducted under these contracts heavily involve school people at the district and state level, funding agencies such as the State legislature or the Federal Government, and one of the most pluralistic societies in the nation.

Findings

Guba and Lincoln strongly advocate the serious inclusion of stakeholders in the process. In a number of cases, legitimate stakeholders have strongly expressed a desire for specific things such as a set of (evaluation) conclusions and a set of (definite) recommendations. If we are to respect these stakeholders' expressed desires, we would attempt to provide such conclusions; however, such definitiveness would be against the philosophy of fourth-generation evaluation.

An example involving the evaluation of federally funded projects

Often federally funded projects provide the funds for evaluation. There usually are federal requirements imposed on the evaluation, and the production of an agenda for negotiations instead of some "hard" data addressing (definitively) these federal requirements directly could result in serious problems for the project. There are indications that the federal government intends to be even more serious about their evaluation requirements and monitoring of evaluation reports.

All evaluations of Title VII bilingual education projects of federally funded evaluations are to follow the 1986 *Bilingual Education Regulations 34 CFR Part 500*. These regulations, for example, require that the evaluator collect student outcome data on the academic achievement of children who were formerly served in the project as limited English proficient, have exited from the program, and are now in English language classrooms (500.50(b)(3)(i)(B)/(ii)(C)). Data must be collected on changes in the rate of student grade retention, dropout, absenteeism, placement in programs for the gifted and talented, and enrollment in postsecondary education institutions (500.52(c)). The evaluation design "must include a measure of the educational progress of project participants when measured against an appropriate nonproject comparison group." (500.50(b)(1)) Evaluation instruments must be administered "at twelve-month testing intervals."

It is not within the scope of this paper to address the question of whether some of the aforementioned regulations are consistent with sound evaluation practice; however, this example is used to make the point that there are legal requirements that are unlikely to be satisfied by the production of an agenda for negotiation to address the relevant claims, concerns, and issues.

Problem in dealing with legislators

Other problems arose in interactions with legislators, whose availability of time was limited so as to not allow extensive negotiations. In reality, those providing simple, brief, and definitive evaluations were the ones that would get listened to by these lawmakers and allocators of funds. Again we have major stakeholders asking for non-fourth-generation evaluation. Given the nature of the legislators' jobs, their requests seem quite reasonable.

Dilemmas arising in our roles as evaluators/stakeholders

As evaluators who have been more or less convinced that it is desirable to (attempt to) conduct fourth-generation evaluations, we are exposed to a number of dilemmas. How do we, in our role as important stakeholders (on the issue of how to conduct evaluations), deal with the fact that some of our current beliefs about and philosophy of evaluation may be in direct conflict with the methods proposed by Guba and Lincoln?

For example, if we are not convinced by Guba and Lincoln that "neither problems nor their solutions can be generalized from one setting to another," then are we doomed to perpetual unenlightenment accompanied by the ability to conduct only third-generation or older evaluations? Or are we entitled to demand our stakeholder rights to negotiate such a claim emanating from the fourth-generation bible?

What if we dare accept some of the positivism that the authors attack rather viciously at times? They claim that "true believers in positivism" regard constructivists and other relativists as "a notch above con men and snake-oil salesmen." What about partial believers in positivism? Even in older-generation research and evaluation, many have cautioned against dichotomizing unjustifiably.

An internal inconsistency in Guba and Lincoln's write-ups

We also note that the authors, despite their own admonition to the contrary (*"...generalizations are not possible."(p. 36)*), make a major generalization from one setting to another very different one. Despite their efforts to clearly distance their procedure from classical scientific inquiry, Guba and Lincoln write the following implied (over-) generalization: "Mary Hesse (1980) has aptly noted that just as all scientific theories have sooner or later proven to be false, so will every theory that we now entertain."(p. 17). A faithful application of the procedures advocated in *Fourth generation evaluation* would avoid making such a generalization from an observation regarding scientific theories to the arena of non-scientific constructivism.

Finally we found what appear to us to be contradictory statements regarding the potential of a "hermeneutic dialectic process." The authors note that they see the possibility that "a new construction will emerge that is not "better" or "truer" that (*sic.*) its predecessors, but simply more informed and sophisticated than either."(p. 17) It is difficult for us to understand why being more informed in not necessarily better, at least in a layperson sense.

Positive aspects of the approach

On the positive side we found it professionally satisfying to assert that our goal was to enhance negotiations rather than to act as if we were omniscient providers of recommendations. It also seemed that in the fourth-generation evaluation arena, qualitative methods and case studies now had appropriately a much improved stature in the evaluation business.

Educational importance

Guba and Lincoln's arguments are convincing, especially about the inadequacy of prior attempts at conducting evaluation. They have obviously spent a substantial effort in designing and presenting their approach. In essence they have offered a theoretical model that, *if applicable in the real world*, promises to dramatically improve evaluation efforts and thereby improve education in general. The importance of this paper is that it describes some of the first empirical attempts at investigating the feasibility of conducting "fourth-generation evaluation." Unless the underlying theory can be effectively applied to the situations with which those in the field must deal, then it is

doomed to be viewed as another impressive, interesting, academic exercise that is basically not applicable in the real world.

Concluding remarks

Newer generations have not necessarily proven to be better than older ones

As stated at the beginning of this paper, we have been positively (not positivistically) influenced by the eloquent arguments made for fourth-generation evaluation. Because we, like everyone else, are not yet well experienced in conducting such evaluations, we have had to evaluate fourth-generation evaluation using essentially third- and older generation evaluation procedures.

We have found serious dilemmas; however, we are willing to continue to pursue the matter further. At the same time we humbly remind ourselves and others that when we look at things like our environment, automobiles, art, music, and literature, we see that the products of more recent generations are not necessarily improvements over those of earlier generations.

References

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.

Hesse, M. (1980). *Revolutions and reconstructions in the philosophy of science*. Bloomington: Indiana University Press.

U.S. Government. (1986, June). Rules and regulations. *Federal Register*, *51*(118), pp. 22425-22426.